

cdna@resgen.com  
DNA Sequencing and analyses performed by National Institutes of Health Intramural Sequencing Center (NISC).

Email: estewatson.wustl.edu  
This clone is available royalty-free through LNL ; contact the  
INRGE Consortium (info@image.lnl.gov) for further information.  
Insert length: 1238 Std Error: 0.00  
Seq primer: mob.REGA+ET  
high quality sequence stop: 64.

FEATURES	FEATURES	SOURCE
High quality sequence	Location/Qualifiers	
1	1	173

```
1. .232
  organism="Xenopus laevis"
  mol_type="mRNA"
  db_xref="taxon:8355"
  clone="PBX0056G11"
  sex="female"
  tissue_type="unfertilized egg"
  cell_type="unfertilized egg"
  dev_stage="unfertilized egg"
  lab_host="DH10B"
  clone_lib="Blackshear/Soares n
  library"
```

library" /note="Vector: pT73-Pac; Site\_1: EcoRI; Site\_2: NotI; PolyA-selected mRNA was prepared from unfertilized Xenopus laevis eggs. The library was constructed in the vector pT73-Pac as described in Bonaldo, M.F., Lennon, G. and Soares, M.B. 'Normalization and subtraction: two approaches to facilitate gene discovery', Genome Research 6:791-806, 1996. The first strand synthesis used a NotI-dT18 primer: double stranded cDNAs were ligated to EcoRI adapters, digested with NotI, and directionally cloned into the NotI and EcoRI-digested pT73-Pac vector. The library contained approximately  $7.2 \times 10^5$  clones."

```

71 a      49 c      65 g      47 t
      100.0%; Score 15; DB 9; Length 232;
      al Similarity 100.0%; Pred. No. 2.1e+02;
      0.0; Mismatches 0; Indels 0; Gaps 0;

```

1 TCCATGGTGCTCACT 15  
|||||  
67 TCCATGGTGCTCACT 81

[illegible]

hereditary multiple exostoses gene 2 (Erx2), human sequence  
AA360512  
AA360512.1 GI:2012902  
EST.  
Homo sapiens (human)

Homo sapiens  
Eumaloryza, Metazoa: Chordata: Craniata: Vertebrata; Euteleostomi;  
Mammalia; Eutheria: Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 268)  
Adams, M. D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult  
C. A., White

C.J.J., Lee, N.H., Kirkness, E.F., Weinstein, K.G., Gockayne, S.,  
O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A.,  
Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald,  
L.M., Fitzhugh, W.M., Frichman, J.L., Geodhagen, N.S., Glodex, A.,  
Kelley, J.M.,

Gnehm, C.L., Hanna, M.C., Hedblom, E., Huikuri, A.L., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Patanques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dinke, D., Feng, D.-F., Ferlie, A., Fischer, C., Hastings, G.A., He, W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Meissner, P.S., Olsen, H., Raymond, D., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillon, P.J., Fann

Fax: 314 286 1810  
Email: estwatson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium ([info@image.lnl.gov](mailto:info@image.lnl.gov)) for further information.  
Insert Length: 1238 Std Error: 0.00  
Seq primer: mob.REGA+ET  
Seq quality sequence stop: 64.

High quality sequence staff	
Location/Qualifiers	
1	173

```
Location/Qualifiers
1. l73
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="GDB:1253939"
  /db_xref="taxon:9608"
  /clone="IMAGE:310523"
  /tissue_type="senescent fibroblast"
  /lab_host="DH10B (ampicillin resistant)"
  /clone_lib="Soares senescent_fibroblasts_NHMF"
  /note="Vector: pT73P (Pharmacia) with a modified
polylinker V-TPPE; phagemid; Site_1: Not I; Site_2: Eco RI;
Site_3: Not I; Site_4: ClaI; Site_5: XbaI; Site_6: SalI; Site_7: PstI; Site_8: KpnI; Site_9: SmaI; Site_10: BamHI; Site_11: XhoI; Site_12: SpeI; Site_13: AclI; Site_14: BspEI; Site_15: BspEII; Site_16: BspEIII; Site_17: BspEIV; Site_18: BspEV; Site_19: BspFVI; Site_20: BspFI; Site_21: BspFII; Site_22: BspFIII; Site_23: BspFIV; Site_24: BspFV; Site_25: BspFVI; Site_26: BspFVII; Site_27: BspFVIII; Site_28: BspFIIX; Site_29: BspFXI; Site_30: BspFXII; Site_31: BspFXIII; Site_32: BspFXIV; Site_33: BspFXV; Site_34: BspFXVI; Site_35: BspFXVII; Site_36: BspFXVIII; Site_37: BspFXIX; Site_38: BspFXXX; Site_39: BspFXXI; Site_40: BspFXXII; Site_41: BspFXXIII; Site_42: BspFXXIV; Site_43: BspFXXV; Site_44: BspFXXVI; Site_45: BspFXXVII; Site_46: BspFXXVIII; Site_47: BspFXXIX; Site_48: BspFXXX; Site_49: BspFXXXI; Site_50: BspFXXXII; Site_51: BspFXXXIII; Site_52: BspFXXXIV; Site_53: BspFXXXV; Site_54: BspFXXXVI; Site_55: BspFXXXVII; Site_56: BspFXXXVIII; Site_57: BspFXXXIX; Site_58: BspFXXXX; Site_59: BspFXXXXI; Site_60: BspFXXXXII; Site_61: BspFXXXXIII; Site_62: BspFXXXXIV; Site_63: BspFXXXXV; Site_64: BspFXXXXVI; Site_65: BspFXXXXVII; Site_66: BspFXXXXVIII; Site_67: BspFXXXXIX; Site_68: BspFXXXXX; Site_69: BspFXXXXXI; Site_70: BspFXXXXXII; Site_71: BspFXXXXXIII; Site_72: BspFXXXXXIV; Site_73: BspFXXXXXV; Site_74: BspFXXXXXVI; Site_75: BspFXXXXXVII; Site_76: BspFXXXXXVIII; Site_77: BspFXXXXXIX; Site_78: BspFXXXXXX; Site_79: BspFXXXXXXI; Site_80: BspFXXXXXXII; Site_81: BspFXXXXXXIII; Site_82: BspFXXXXXXIV; Site_83: BspFXXXXXXV; Site_84: BspFXXXXXXVI; Site_85: BspFXXXXXXVII; Site_86: BspFXXXXXXVIII; Site_87: BspFXXXXXXIX; Site_88: BspFXXXXXXX; Site_89: BspFXXXXXXXI; Site_90: BspFXXXXXXXII; Site_91: BspFXXXXXXXIII; Site_92: BspFXXXXXXXIV; Site_93: BspFXXXXXXXV; Site_94: BspFXXXXXXXVI; Site_95: BspFXXXXXXXVII; Site_96: BspFXXXXXXXVIII; Site_97: BspFXXXXXXXIX; Site_98: BspFXXXXXXXX; Site_99: BspFXXXXXXXXI; Site_100: BspFXXXXXXXXII; Site_101: BspFXXXXXXXXIII; Site_102: BspFXXXXXXXXIV; Site_103: BspFXXXXXXXXV; Site_104: BspFXXXXXXXXVI; Site_105: BspFXXXXXXXXVII; Site_106: BspFXXXXXXXXVIII; Site_107: BspFXXXXXXXXIX; Site_108: BspFXXXXXXXXI; Site_109: BspFXXXXXXXXII; Site_110: BspFXXXXXXXXIII; Site_111: BspFXXXXXXXXIV; Site_112: BspFXXXXXXXXV; Site_113: BspFXXXXXXXXVI; Site_114: BspFXXXXXXXXVII; Site_115: BspFXXXXXXXXVIII; Site_116: BspFXXXXXXXXIX; Site_117: BspFXXXXXXXXI; Site_118: BspFXXXXXXXXII; Site_119: BspFXXXXXXXXIII; Site_120: BspFXXXXXXXXIV; Site_121: BspFXXXXXXXXV; Site_122: BspFXXXXXXXXVI; Site_123: BspFXXXXXXXXVII; Site_124: BspFXXXXXXXXVIII; Site_125: BspFXXXXXXXXIX; Site_126: BspFXXXXXXXXI; Site_127: BspFXXXXXXXXII; Site_128: BspFXXXXXXXXIII; Site_129: BspFXXXXXXXXIV; Site_130: BspFXXXXXXXXV; Site_131: BspFXXXXXXXXVI; Site_132: BspFXXXXXXXXVII; Site_133: BspFXXXXXXXXVIII; Site_134: BspFXXXXXXXXIX; Site_135: BspFXXXXXXXXI; Site_136: BspFXXXXXXXXII; Site_137: BspFXXXXXXXXIII; Site_138: BspFXXXXXXXXIV; Site_139: BspFXXXXXXXXV; Site_140: BspFXXXXXXXXVI; Site_141: BspFXXXXXXXXVII; Site_142: BspFXXXXXXXXVIII; Site_143: BspFXXXXXXXXIX; Site_144: BspFXXXXXXXXI; Site_145: BspFXXXXXXXXII; Site_146: BspFXXXXXXXXIII; Site_147: BspFXXXXXXXXIV; Site_148: BspFXXXXXXXXV; Site_149: BspFXXXXXXXXVI; Site_150: BspFXXXXXXXXVII; Site_151: BspFXXXXXXXXVIII; Site_152: BspFXXXXXXXXIX; Site_153: BspFXXXXXXXXI; Site_154: BspFXXXXXXXXII; Site_155: BspFXXXXXXXXIII; Site_156: BspFXXXXXXXXIV; Site_157: BspFXXXXXXXXV; Site_158: BspFXXXXXXXXVI; Site_159: BspFXXXXXXXXVII; Site_160: BspFXXXXXXXXVIII; Site_161: BspFXXXXXXXXIX; Site_162: BspFXXXXXXXXI; Site_163: BspFXXXXXXXXII; Site_164: BspFXXXXXXXXIII; Site_165: BspFXXXXXXXXIV; Site_166: BspFXXXXXXXXV; Site_167: BspFXXXXXXXXVI; Site_168: BspFXXXXXXXXVII; Site_169: BspFXXXXXXXXVIII; Site_170: BspFXXXXXXXXIX; Site_171: BspFXXXXXXXXI; Site_172: BspFXXXXXXXXII; Site_173: BspFXXXXXXXXIII; Site_174: BspFXXXXXXXXIV; Site_175: BspFXXXXXXXXV; Site_176: BspFXXXXXXXXVI; Site_177: BspFXXXXXXXXVII; Site_178: BspFXXXXXXXXVIII; Site_179: BspFXXXXXXXXIX; Site_180: BspFXXXXXXXXI; Site_181: BspFXXXXXXXXII; Site_182: BspFXXXXXXXXIII; Site_183: BspFXXXXXXXXIV; Site_184: BspFXXXXXXXXV; Site_185: BspFXXXXXXXXVI; Site_186: BspFXXXXXXXXVII; Site_187: BspFXXXXXXXXVIII; Site_188: BspFXXXXXXXXIX; Site_189: BspFXXXXXXXXI; Site_190: BspFXXXXXXXXII; Site_191: BspFXXXXXXXXIII; Site_192: BspFXXXXXXXXIV; Site_193: BspFXXXXXXXXV; Site_194: BspFXXXXXXXXVI; Site_195: BspFXXXXXXXXVII; Site_196: BspFXXXXXXXXVIII; Site_197: BspFXXXXXXXXIX; Site_198: BspFXXXXXXXXI; Site_199: BspFXXXXXXXXII; Site_200: BspFXXXXXXXXIII; Site_201: BspFXXXXXXXXIV; Site_202: BspFXXXXXXXXV; Site_203: BspFXXXXXXXXVI; Site_204: BspFXXXXXXXXVII; Site_205: BspFXXXXXXXXVIII; Site_206: BspFXXXXXXXXIX; Site_207: BspFXXXXXXXXI; Site_208: BspFXXXXXXXXII; Site_209: BspFXXXXXXXXIII; Site_210: BspFXXXXXXXXIV; Site_211: BspFXXXXXXXXV; Site_212: BspFXXXXXXXXVI; Site_213: BspFXXXXXXXXVII; Site_214: BspFXXXXXXXXVIII; Site_215: BspFXXXXXXXXIX; Site_216: BspFXXXXXXXXI; Site_217: BspFXXXXXXXXII; Site_218: BspFXXXXXXXXIII; Site_219: BspFXXXXXXXXIV; Site_220: BspFXXXXXXXXV; Site_221: BspFXXXXXXXXVI; Site_222: BspFXXXXXXXXVII; Site_223: BspFXXXXXXXXVIII; Site_224: BspFXXXXXXXXIX; Site_225: BspFXXXXXXXXI; Site_226: BspFXXXXXXXXII; Site_227: BspFXXXXXXXXIII; Site_228: BspFXXXXXXXXIV; Site_229: BspFXXXXXXXXV; Site_230: BspFXXXXXXXXVI; Site_231: BspFXXXXXXXXVII; Site_232: BspFXXXXXXXXVIII; Site_233: BspFXXXXXXXXIX; Site_234: BspFXXXXXXXXI; Site_235: BspFXXXXXXXXII; Site_236: BspFXXXXXXXXIII; Site_237: BspFXXXXXXXXIV; Site_238: BspFXXXXXXXXV; Site_239: BspFXXXXXXXXVI; Site_240: BspFXXXXXXXXVII; Site_241: BspFXXXXXXXXVIII; Site_242: BspFXXXXXXXXIX; Site_243: BspFXXXXXXXXI; Site_244: BspFXXXXXXXXII; Site_245: BspFXXXXXXXXIII; Site_246: BspFXXXXXXXXIV; Site_247: BspFXXXXXXXXV; Site_248: BspFXXXXXXXXVI; Site_249: BspFXXXXXXXXVII; Site_250: BspFXXXXXXXXVIII; Site_251: BspFXXXXXXXXIX; Site_252: BspFXXXXXXXXI; Site_253: BspFXXXXXXXXII; Site_254: BspFXXXXXXXXIII; Site_255: BspFXXXXXXXXIV; Site_256: BspFXXXXXXXXV; Site_257: BspFXXXXXXXXVI; Site_258: BspFXXXXXXXXVII; Site_259: BspFXXXXXXXXVIII; Site_260: BspFXXXXXXXXIX; Site_261: BspFXXXXXXXXI; Site_262: BspFXXXXXXXXII; Site_263: BspFXXXXXXXXIII; Site_264: BspFXXXXXXXXIV; Site_265: BspFXXXXXXXXV; Site_266: BspFXXXXXXXXVI; Site_267: BspFXXXXXXXXVII; Site_268: BspFXXXXXXXXVIII; Site_269: BspFXXXXXXXXIX; Site_270: BspFXXXXXXXXI; Site_271: BspFXXXXXXXXII; Site_272: BspFXXXXXXXXIII; Site_273: BspFXXXXXXXXIV; Site_274: BspFXXXXXXXXV; Site_275: BspFXXXXXXXXVI; Site_276: BspFXXXXXXXXVII; Site_277: BspFXXXXXXXXVIII; Site_278: BspFXXXXXXXXIX; Site_279: BspFXXXXXXXXI; Site_280: BspFXXXXXXXXII; Site_281: BspFXXXXXXXXIII; Site_282: BspFXXXXXXXXIV; Site_283: BspFXXXXXXXXV; Site_284: BspFXXXXXXXXVI; Site_285: BspFXXXXXXXXVII; Site_286: BspFXXXXXXXXVIII; Site_287: BspFXXXXXXXXIX; Site_288: BspFXXXXXXXXI; Site_289: BspFXXXXXXXXII; Site_290: BspFXXXXXXXXIII; Site_291: BspFXXXXXXXXIV; Site_292: BspFXXXXXXXXV; Site_293: BspFXXXXXXXXVI; Site_294: BspFXXXXXXXXVII; Site_295: BspFXXXXXXXXVIII; Site_296: BspFXXXXXXXXIX; Site_297: BspFXXXXXXXXI; Site_298: BspFXXXXXXXXII; Site_299: BspFXXXXXXXXIII; Site_300: BspFXXXXXXXXIV; Site_301: BspFXXXXXXXXV; Site_302: BspFXXXXXXXXVI; Site_303: BspFXXXXXXXXVII; Site_304: BspFXXXXXXXXVIII; Site_305: BspFXXXXXXXXIX; Site_306: BspFXXXXXXXXI; Site_307: BspFXXXXXXXXII; Site_308: BspFXXXXXXXXIII; Site_309: BspFXXXXXXXXIV; Site_310: BspFXXXXXXXXV; Site_311: BspFXXXXXXXXVI; Site_312: BspFXXXXXXXXVII; Site_313: BspFXXXXXXXXVIII; Site_314: BspFXXXXXXXXIX; Site_315: BspFXXXXXXXXI; Site_316: BspFXXXXXXXXII; Site_317: BspFXXXXXXXXIII; Site_318: BspFXXXXXXXXIV; Site_319: BspFXXXXXXXXV; Site_320: BspFXXXXXXXXVI; Site_321: BspFXXXXXXXXVII; Site_322: BspFXXXXXXXXVIII; Site_
```

polylinker V-type; phage DNA was primed with a Not I - oligo(dT) primer [5],  
 TGGTTACCAATCTGAAGTGGGAGCGCCGCATTGTTCCTTTTTTTT 3'}, RI  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified pT7T3 vector  
 (Pharmacia). Library went through one round of  
 normalization to a Cot = 5. Library constructed by Bento  
 Soares and M.Fatima Bonaído."

```

each 100.0%; score 10.1284;
15 Similarity 100.0%; Pred. No. 1.9e+02;
15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCCATGGTGCTACT 15
|||||

```

bl56gll.wl Blackshear/Soares normalized message  
laevis cDNA clone PX0056G11 5', mRNA sequence.  
AW637280  
AW637280.1 GI:7394388

EST.  
Xenopus laevis (African clawed frog)  
Xenopus laevis  
4 Xenopus laevis  
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;

xenopodinae: xenopus.  
1 (bases 1 to 232)  
Blackshear, P.J., Lai, W.S., Thorn, J.M., Kennington, E.A., Staffa, N.G.  
Jr., Moore, D.T., Bouffard, G.G., Beckstrom-Sternberg, S.M., Touchman  
J.W., Ronaldo, M.F. and Soares, M.B.  
1989. *in press* protect. interim analysis of the

The NIEHS Xenopus maternal ESI project  
first 13,879 ESTs from unfertilized eggs  
Gene 267 (1), 71-87 (2001)  
21211403

1131155/  
Contact: Perry J. Blackshear  
Office of Clinical Research and Laboratory of Signal Transduction  
National Institute of Environmental Health Sciences  
A2-05 NIEHS, 101 Alexander Drive, Research Triangle Park, NC 27709  
USA  
Tel: 919 541-4899  
Fax: 919 541-4571  
Email: black009@niehs.nih.gov  
\*Clone is available through Research Genetics, Inc., 2130 Memorial  
Avenue, Huntsville, AL 35901

Day, Huntsville, AL 35894-0001, fax 256-536-9016 att:cdna, email  
cdna-536-9016 ext:cdna.



M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.  
Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence  
Nature 377 (6547 Suppl), 3-174 (1995)  
96026280  
MEDLINE  
PUBMED  
COMMENT  
Other ESTs: THC194116  
7566098  
Contact: Kerlavage, AR  
Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423  
Email: arkerlavet@tigr.org  
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)  
Seq primer: M13 Reverse.  
Location/Qualifiers

## FEATURES

source  
1..268  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="ATCC (inhost):164695"  
/db\_xref="taxon:9606"  
/cell\_type="T-lymphocyte"  
/clone\_lib="T-cell lymphoma"  
/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI"

BASE COUNT 67 a 46 c 77 g 76 t 2 others  
ORIGIN

Query Match 100.0%; Score 15; DB 9; Length 268;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 TCCATGGTGCTCACT 15  
|||||  
Db 248 TCCATGGTGCTCACT 262

RESULT 4  
BE132222  
LOCUS BE132222 269 bp mRNA linear EST 29-JUN-2000  
DEFINITION db43901.y1 Blackshear/Soares normalized xenopus egg library Xenopus laevis cDNA clone IMAGE:3301296 5', mRNA sequence.  
ACCESSION BE132222  
VERSION BE132222.1 GI:8579587  
KEYWORDS EST.  
SOURCE Xenopus laevis (African clawed frog)  
ORGANISM Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae; Xenopus.

REFERENCE 1 (bases 1 to 269)  
AUTHORS Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Willson, R.

TITLE WashU Xenopus EST project, 1999  
JOURNAL Unpublished  
COMMENT Contact: Sandy Clifton, Ph.D.  
WashU Xenopus EST project, 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@watson.wustl.edu  
Library constructed by Bento Soares and M. Fatima Bonaldo (University of Iowa). DNA Sequencing by: Washington University Genome Sequencing Center.  
Clone Distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/LLNL at: [info@image.llnl.gov](mailto:info@image.llnl.gov)

Seq primer: -40RP from Gibco  
High quality sequence stop: 268.

## FEATURES

source  
1..269  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="IMAGE:3301296"  
/sex="female"  
/tissue\_type="unfertilized egg"  
/cell\_type="unfertilized egg"  
/dev\_stage="unfertilized egg"  
/lab\_host="DH10B"  
/clone\_lib="Blackshear/Soares normalized Xenopus egg library"

/note="Vector: pT7T3-Pac; Site\_1: EcoRI; Site\_2: NotI;  
Polya-selected mRNA was prepared from unfertilized Xenopus laevis eggs. The library was constructed in the vector pT7T3-Pac as described in Bonaldo, M.F., Lennon, G. and Soares, M.B. 'Normalization and subtraction: two approaches to facilitate gene discovery', Genome Research 6:791-806, 1996. The first strand synthesis used a NotI-dT18 primer; double stranded cDNAs were ligated to EcoRI adapters, digested with NotI, and directionally cloned into the NotI and EcoRI-digested pT7T3-Pac vector. The library contained approximately 7.2 x 10<sup>5</sup> recombinants, with average insert sizes of 1-1.5 kb."

BASE COUNT 79 a 61 c 55 t 1 others  
ORIGIN

Query Match 100.0%; Score 15; DB 10; Length 269;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 TCCATGGTGCTCACT 15  
|||||  
Db 196 TCCATGGTGCTCACT 210

RESULT 5  
BF822715  
LOCUS BF822715 321 bp mRNA linear EST 13-JAN-2001  
DEFINITION CM3-RT00006-091200-545-c12 RT00006 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF822715  
VERSION BF822715.1 GI:12162929  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 321)  
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bal, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
PUBMED 10737800

COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&t2=CM3-RT0006->



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 11, 2003, 10:13:13 ; Search time 2294 Seconds  
(without alignments)  
158.922 Million cell updates/sec

Title: US-09-716-320-3

Perfect score: 15

Sequence: 1 tccatggtgcact 15

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estlin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_estc:\*

9: gb\_estli:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_iman:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	100.0	173	14 W31091	W31091 zb86h07.r1
2	15	100.0	232	9 AW637280	AW637280 bl56g11.w
3	15	100.0	268	9 AA360512	AA360512 EST69742
4	15	100.0	269	10 BE132222	BE132222 db43g01.y

5	15	100.0	321	10 BF822715	BF822715 CM3-RT000
6	15	100.0	322	9 A1008593	A1008593 EST203044
7	15	100.0	329	12 BI042474	BI042474 BC5-OT009
8	15	100.0	357	13 BY400016	BY400016 BY400016
9	15	100.0	379	9 AA085493	AA085493 zn44c07.f
10	15	100.0	381	10 BG121876	BG121876 602351302
11	15	100.0	389	10 AW886682	AW886682 RCI-OT008
12	15	100.0	393	12 BI057614	BI057614 MRO-GN017
13	15	100.0	398	28 BH096821	BH096821 RPCI-24-2
14	15	100.0	400	28 AQ296272	AQ296272 HS-2166_B
15	15	100.0	432	14 R32275	R32275 yh68e03.r1
16	15	100.0	444	13 BY556459	BY556459 BY556459
17	15	100.0	459	9 AW636779	AW636779 bl49g11.w
18	15	100.0	463	12 BM708050	BM708050 UI-E-C11-
19	15	100.0	470	9 AU243882	AU243882 AU243882
20	15	100.0	474	14 N31328	N31328 yx54b06.r1
21	15	100.0	484	9 AA312237	AA312237 EST183053
22	15	100.0	512	9 AA452475	AA452475 zx31b07.f
23	15	100.0	513	29 BZ873773	BZ873773 CH240_292
24	15	100.0	514	10 BE018291	BE018291 bb78c02.y
25	15	100.0	521	12 BM823695	BM823695 K-EST0094
26	15	100.0	523	10 BE971148	BE971148 601651393
27	15	100.0	524	10 BF875537	BF875537 QV3-ET010
28	15	100.0	577	9 AL636508	AL636508 AL636508
29	15	100.0	594	13 BQ308056	BQ308056 MRO-BT550
30	15	100.0	604	9 AV714029	AV714029 AV714029
31	15	100.0	606	12 BJ029020	BJ029020 BJ029020
32	15	100.0	616	10 BE867032	BE867032 601442866
33	15	100.0	666	12 BI891968	BI891968 EF637-3-0
34	15	100.0	672	12 BM724884	BM724884 UI-E-E01-
35	15	100.0	679	29 AG103671	AG103671 Pan trogl
36	15	100.0	690	28 AZ733709	AZ733709 RPCI-24-1
37	15	100.0	722	12 BI871769	BI871769 603394022
38	15	100.0	731	10 BE793514	BE793514 601588859
39	15	100.0	743	10 BF792458	BF792458 602253541
40	15	100.0	747	14 CD511710	CD511710 AGENCOURT
41	15	100.0	772	10 BE790912	BE790912 601582349
42	15	100.0	776	9 AU132253	AU132253 AU132253
43	15	100.0	786	10 BG435979	BG435979 602508347
44	15	100.0	789	10 BE799238	BE799238 601586816
45	15	100.0	816	10 BG683112	BG683112 602651791

ALIGNMENTS

RESULT 1

W31091

LOCUS

DEFINITION

W31091

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

W31091 173 bp mRNA linear EST 20-AUG-1996  
zb86h07.r1 Soares\_senescent\_fibroblasts\_NBHSF Homo sapiens cDNA  
clone IMAGE:310525 5', mRNA sequence.

W31091.1 GI:1312081

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 173)

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,

Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins

, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore

, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,

Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E.,

Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M.

Generation and analysis of 280,000 human expressed sequence tags

Genome Res. 6 (9), 807-828 (1996)

97044478

8889549

CONTACT: Willson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810  
 Email: estewatson.wustl.edu  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 1238 Std Error: 0.00  
 Seq primer: mob.REGA-ET  
 High quality sequence stop: 64.  
 Location/Qualifiers

# FEATURES

1..173

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:1253939"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:310525"  
 /tissue\_type="senescent fibroblast"  
 /lab\_host="PH10B (ampicillin resistant)"  
 /clone\_lib="Soares\_senescent\_fibroblasts\_NBHSF"  
 /note="vector: pT73D (Pharmacia) with a modified  
 polylinker V-TYPE: phagemid; Site\_1: Not I; Site\_2: Eco RI  
 ; 1st strand cDNA was primed with a Not I - oligo(dT)  
 primer [5'  
 TGTACCAATCTGAAGTGGGAGCGCGCATTTTTTTTTTTTTTTT 3'],  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified pT73 vector  
 (Pharmacia). Library went through one round of  
 normalization to a Cot = 5. Library constructed by Bento  
 Soares and M.Fatima Bonaldo."  
 Soares 43 a 33 c 48 g 49 t

# BASE COUNT

Query Match 100.0%; Score 15; DB 14; Length 173;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCATGGTGCTCACT 15

|||||

D5 93 TCCATGGTGCTCACT 107

# RESULT 2

AW637280

LOCUS

DEFINITION B156g11.w1 Blackshear/Soares normalized Xenopus egg library Xenopus  
 laevis cDNA clone PBX0056G11 5', mRNA sequence.

AW637280

AW637280.1 GI:7394388

EST.

Xenopus laevis (African clawed frog)

Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

Xenopodinae; Xenopus.

1 (bases 1 to 232)

Blackshear,P.J., Lai,W.S., Thorn,J.M., Kennington,E.A., Staffa,N.G.

Jr., Moore,D.T., Bouffard,G.G., Beckstrom-Sternberg,S.M., Touchman

J.W., Bonaldo,M.F. and Soares,M.B.

The NIEHS Xenopus maternal EST project: Interim analysis of the

first 13,879 ESTs from unfertilized eggs

Gene 267 (1), 71-87 (2001)

21211403

11311557

COMMENT

Contact: Perry J. Blackshear

Office of Clinical Research and Laboratory of Signal Transduction

National Institute of Environmental Health Sciences

A2-05 NIEHS, 101 Alexander Drive, Research Triangle Park, NC 27709,

USA

Tel: 919 541-4899

Fax: 919 541-4571

Email: black009@niehs.nih.gov

Clone is available through Research Genetics, Inc., 2130 Memorial

Parkway, Huntsville, AL 35901

phone 800-533-4363 ext.cdna, fax 256-536-9016 att.cdna, email

cdna@resgen.com

DNA Sequencing and analyses performed by National Institutes of  
 Health Intramural Sequencing Center (NISC).

PCR Primers

FORWARD: TGTAAACAGCGCCAGT

BACKWARD: CAGAAACAGCTATGACC

Plate: 0056 row: G column: 11

Seq primer: T7 primer.

# FEATURES

source

1..232

/organism="Xenopus laevis"

/mol\_type="mRNA"

/db\_xref="taxon:8355"

/clone="PBX0056G11"

/sex="female"

/tissue\_type="unfertilized egg"

/cell\_type="unfertilized egg"

/dev\_stage="unfertilized egg"

/lab\_host="BH10B"

/clone\_lib="Blackshear/Soares normalized Xenopus egg  
 library"

/note="Vector: pT73-Pac; Site\_1: EcoRI; Site\_2: NotI;  
 PolyA-selected mRNA was prepared from unfertilized Xenopus  
 laevis eggs. The library was constructed in the vector  
 pT73-Pac as described in Bonaldo, M.F., Lennon, G. and  
 Soares, M.B. 'Normalization and subtraction: two  
 approaches to facilitate gene discovery', Genome Research  
 6:791-806, 1996. The first strand synthesis used a  
 NotI-dT18 primer; double stranded cDNAs were ligated to  
 EcoRI adapters, digested with NotI, and directionally  
 cloned into the NotI and EcoRI-digested pT73-Pac vector.  
 The library contained approximately 7.2 X 10<sup>5</sup>  
 recombinants, with average insert sizes of 1-1.5 kb."

BASE COUNT 71 a 49 c 65 g 47 t

# ORIGIN

Query Match 100.0%; Score 15; DB 9; Length 232;

Best Local Similarity 100.0%; Pred. No. 2.1e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCATGGTGCTCACT 15

|||||

D5 67 TCCATGGTGCTCACT 81

# RESULT 3

AA360512

LOCUS

DEFINITION

AA360512

AA360512

AA360512.1 GI:2012902

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 268)

ADAMS, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult

C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White

O., Sutton, G., Blake, J.A., Brandon, R.C., Men-Wei, C., Clayton, R.A.,

Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald

L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S., Glodek, A.,

Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M.,

Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,

Moreno-Palances, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,

Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R.,

Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,

Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,

Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W.

Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L.,

Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L.,

Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon

268 bp mRNA linear EST 21-APR-1997

EST69742 T-cell lymphoma Homo sapiens cDNA 5' end similar to

hereditary multiple exostoses gene 2 (EXT2), mRNA sequence.

AA360512

AA360512.1 GI:2012902

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 268)

ADAMS, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult

C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White

O., Sutton, G., Blake, J.A., Brandon, R.C., Men-Wei, C., Clayton, R.A.,

Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald

L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S., Glodek, A.,

Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M.,

Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,

Moreno-Palances, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,

Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R.,

Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,

Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,

Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W.

Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L.,

Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L.,

Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon

M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.  
Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence  
Nature 377 (5547 Suppl), 3-174 (1995)  
96026280  
7565098  
PUBMED  
COMMENT

Other\_ESTs: THC194116  
Contact: Kerlavage, AR  
Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423

Email: arkerlav@tigr.org  
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)  
Seq primer: M13 Reverse.

#### FEATURES

Location/Qualifiers  
1..268  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="ATCC (inhost):164695"  
/db\_xref="taxon:9608"  
/cell\_type="T-lymphocyte"  
/clone\_lib="T-cell lymphoma"  
/note="vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI"

BASE COUNT 67 a 46 c 77 g 76 t 2 others  
ORIGIN  
Query Match 100.0%; Score 15; DB 9; Length 268;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCATGGTGCTCACT 15  
|||||  
Db 248 TCCATGGTGCTCACT 262

RESULT 4  
BF132222  
LOCUS BF132222 269 bp mRNA linear EST 29-JUN-2000  
DEFINITION db43901.y1 Blackshear/Soares normalized xenopus egg library Xenopus laevis cDNA clone IMAGE:3301296 5', mRNA sequence.  
ACCESSION BF132222  
VERSION BF132222.1 GI:8579587  
KEYWORDS EST.  
SOURCE Xenopus laevis (African clawed frog)

ORGANISM  
Xenopus laevis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae; Xenopus.  
1 (bases 1 to 269)  
Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Willson, R.  
WashU Xenopus EST project, 1999  
Unpublished  
Contact: Sandy Clifton, Ph.D.  
WashU Xenopus EST project, 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu

Library constructed by Bento Soares and M. Fatima Bonaldo (University of Iowa). DNA Sequencing by: Washington University Genome Sequencing Center.  
Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/LLNL at: [info@image.llnl.gov](mailto:info@image.llnl.gov)

Seq primer: -40RP from Gibco  
High quality sequence stop: 268.  
Location/Qualifiers  
1..269  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="IMAGE:3301296"  
/sex="female"  
/tissue\_type="unfertilized egg"  
/cell\_type="unfertilized egg"  
/dev\_stage="unfertilized egg"  
/lab\_host="DH10B"  
/clone\_lib="Blackshear/Soares normalized xenopus egg library"

/note="vector: pT73-Pac; Site\_1: EcoRI; Site\_2: NotI; polyA-selected mRNA was prepared from unfertilized Xenopus laevis eggs. The library was constructed in the vector pT73-Pac as described in Bonaldo, M.F., Lennon, G. and Soares, M.B. 'Normalization and subtraction: two approaches to facilitate gene discovery', Genome Research 6:791-806, 1996. The first strand synthesis used a NotI-dn18 primer; double stranded cDNAs were ligated to EcoRI adapters, digested with NotI, and directionally cloned into the NotI and EcoRI-digested pT73-Pac vector. The library contained approximately 7.2 x 10<sup>5</sup> recombinants, with average insert sizes of 1-1.5 kb."  
BASE COUNT 79 a 61 c 73 g 55 t 1 others  
ORIGIN

Query Match 100.0%; Score 15; DB 10; Length 269;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCCATGGTGCTCACT 15  
|||||  
Db 196 TCCATGGTGCTCACT.210

RESULT 5  
BF822715  
LOCUS BF822715 321 bp mRNA linear EST 13-JAN-2001  
DEFINITION CM3-RF0006-091200-545-cl2 RT0006 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF822715  
VERSION BF822715.1 GI:12162929  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 321)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, C.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

TITLE  
PROC. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&t=CM3-RF0006->

091200-545-cl2&t3-2000-12-09&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 3  
High quality sequence stop: 321.  
Location/Qualifiers

# FEATURES

source  
1..321  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="RT0006"  
/note="Organ: kidney\_tumor; Vector: puc18; Site\_1: SmaI;  
Site\_2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."  
88 a 53 c 88 g 82 t  
BASE COUNT  
ORIGIN  
Query Match 100.0%; Score 15; DB 10; Length 321;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TCCATGGTGTCTACT 15  
|||||  
Db 110 TCCATGGTGTCTACT 124

RESULT 6  
AI008593/c  
LOCUS  
DEFINITION EST203044 Normalized rat embryo, Bento Soares Rattus sp. cDNA clone  
REMA248 3' end, mRNA sequence.

ACCESSION AI008593  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Rattus sp.  
Rattus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus

REFERENCE 1 (bases 1 to 322)

AUTHORS Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,  
Kerlavage, A.R. and Adams, M.D.

TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat  
Gene Index

JOURNAL Unpublished

COMMENT Contact: Lee, NH  
The Institute for Genomic Research  
9712, Medical Center Drive, Rockville, MD 20850, USA  
Tel: (301)-838-3529  
Fax: (301)-838-0208  
Email: nhlee@tigr.org

Seq primer: M13-21.

FEATURES  
source  
1..322  
Location/Qualifiers

/organism="Rattus sp."  
/mol\_type="mRNA"  
/db\_xref="ATCC (inhost):2016778"  
/db\_xref="taxon:10118"  
/clone="REMA248"  
/dev\_stage="embryo 8, 12, 18 dpc"  
/clone\_lib="Normalized rat embryo, Bento Soares"  
/note="Vector: pT7r3pac; Site\_1: EcoRI; Site\_2: NotI"

BASE COUNT 63 a 94 c 101 g 64 t  
ORIGIN  
Query Match 100.0%; Score 15; DB 9; Length 322;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TCCATGGTGTCTACT 15  
|||||  
Db 321 TCCATGGTGTCTACT 307

# RESULT 7

BI042474  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE 1 (bases 1 to 329)  
AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC5&t2=RC5-OT0098-  
220101-011-C05&t3=2001-01-22&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 82  
High quality sequence stop: 327.  
Location/Qualifiers

1..329  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="OT0098"

/note="Organ: ovary;  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No. 196  
716 - Ludwig Institute for Cancer Research) profiles  
into the puc 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."

BASE COUNT 87 a 59 c 93 g 89 t 1 others  
ORIGIN

Query Match 100.0%; Score 15; DB 12; Length 329;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TCCATGGTGTCTACT 15  
|||||  
Db 285 TCCATGGTGTCTACT 299

# RESULT 8

BY400016/c  
LOCUS  
DEFINITION

Query Match 100.0%; Score 15; DB 12; Length 329;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

BY400016  
BY400016 RIKEN full-length enriched, CRL-1722 L5178Y-R Mus musculus



CDNA clone I730028H04 3', mRNA sequence.

ACCESSION  
BY400016  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

mus musculus (house mouse)

REFERENCE  
AUTHORS  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. (bases 1 to 357)

Okazaki, Y., Futuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H.

Gojohori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.-Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sakai, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsr.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/

Aizawa, K., Akimura, T., Arakawa, F., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

FEATURES  
source

1. .357

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="DBA/2"

/db\_xref="taxon:10090"

/clone="I730028H04"

/cell\_line="CRL-1722 L5178Y-R"

/clone\_lib="RIKEN full-length enriched, CRL-1722 L5178Y-R"

/clone\_11b="80 g 91 t 1 others

BASE COUNT 106 a 79 c 80 g 91 t

ORIGIN

Query Match 100.0%; Score 15; DB 13; Length 357;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCATGGTGCTCACT 15

|||||

65 TCCATGGTGCTCACT 51

RESULT 9

AA085493

LOCUS

AA085493

DEFINITION

zn44c07.r1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone

IMAGE:550284 5' similar to TR:E229331 E229331 PROTEIN RESPONSIBLE

FOR HEREDITARY MULTIPLE EXOSTOSIS. ; mRNA sequence.

ACCESSION

AA085493

VERSION

AA085493.1 GI:1628769

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 379)

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,

Chissoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.

Generation and analysis of 280,000 human expressed sequence tags

Genome Res. 6 (9), 807-828 (1996)

97044478

889549

COMMENT

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.

Insert Length: 1156 Std Error: 0.00

Seq primer: -28M13 rev2 from Amersham.

Location/Qualifiers

1. .379

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:3928802"

/db\_xref="taxon:9606"

/clone="IMAGE:550284"

/sex="female"

/dev\_stage="HeLa S3 cell line"

/lab\_host="SOLR (kanamycin resistant)"

/clone\_lib="Stratagene HeLa cell s3 937216"

/notes="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. HeLa S3 epithelioid carcinoma cells grown to semi-confluency without induction. Average insert size: 1.5 kb; Uni-ZAP XR Vector. -5' adaptor sequence: 5' GAATTCGGCAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' "

108 a 88 c 92 g 91 t

BASE COUNT



(pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choori.org/bacpac/orderingframe.htm>). BAC end page: [http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
Plate: 298 row: D column: 11  
Seq primer: SP6  
Class: BAC ends.

	Location/Qualifiers
1..398	
/organism="Mus musculus"	
/mol_type="genomic DNA"	
/strain="C57BL/6J"	
/db_xref="taxon:10090"	
/clone="RPCI-24-298D11"	
/sex="Male"	
/cell_type="Spleen/Brain"	
/clone_lib="RPCI-24"	
/note="vector: pTARBAC1; Site_1: BamH1; Site_2: BamH1; RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamH1 sites using MboI partially digested male C57BL/6J DNA."	

```

100.0%; Score 15; DB 28; Length 398;
Similarity 100.0%; Pred. No. 2.5e+02;
15; Conservative 0; Mismatches 0; Indels 0; Gaps 0

1 TCCATGGTGCTCACT 15
|||||
0 TCCATGGTGCTCACT 236

```

AQ296272.1 400 bp UNA linear GSS 13-DEC-1999  
 HS-2166\_B2\_WR\_D05 CIT Approved Human Genomic Sperm Library D Homo  
 sapiens genomic clone Plate=2166 Col=10 Row=H, genomic survey  
 sequence.  
 AQ296272 AQ296272  
 AQ296272.1 GI:4010340  
 GSS.  
 Homo sapiens (human)  
 Homo sapiens  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 400)  
 MaHairs,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and  
 Hood,L.

scanning the human genome  
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
99380589  
10449764  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: thallag@u.washington.edu

```

Plate: 2166   row: H   column: 10
Class: BAC ends
High quality sequence stop: 400.
Location/Qualifiers
    1. .400
       /organism="Homo sapiens"
       /mol_type="genomic DNA"
       /db_xref="taxon:9606"
       /clone="Plate=2166 Col=10 Row=H"
       /sex="male"

```

/clone\_lib="CIT:Approved Human Genomic Sperm Library D"  
/note="Organ: sperm; Vector: pBelOAC11; BAC Clones in  
E-Coli DH10B"

BASE COUNT 99 a 95 c 95 g 111 t

Query Match 100.0%; Score 15; DB 28; Length 400;

Best Local Similarity 100.0%; Pred. NO. 2.5e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCATGGTGCTCACT 15  
|||||  
Db 254 TCCATGGTGCTCACT 268

RESULT 15

R32275

LOCUS

DEFINITION

YH68603.r1 Soares placenta Nb2HP Homo sapiens cDNA clone

IMAGE:134908 5', mRNA sequence.

ACCESSION R32275

VERSION R32275.1

KEYWORDS GI:788118

SOURCE EST.

ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 432)

AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman

,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,

Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston

,R., Williamson,A., Wohlmann,P. and Wilson,R.

THE WashU-Merck EST Project

Unpublished

CONTACT: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert Size: 1377

High quality sequence stops: 383

Source: IMAGE Consortium, LNL

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

Insert Length: 1377 Std Error: 0.00

Seq primer: M13RP1

High quality sequence stop: 383.

FEATURES

Location/Qualifiers

1..432

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:540785"

/db\_xref="taxon:9606"

/clone="IMAGE:134908"

/sex="Female"

/dev\_stage="placenta obtained at birth (full term)"

/lab\_host="DH10B (ampicillin resistant)"

/clone\_lib="Soares placenta Nb2HP"

/note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a

modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer (5'

AACTGAGAAATTCGCGCGAGGATTTTTTTTTTTT 3'),

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT73 vector. Library

went through one round of normalization. Library

constructed by Bento Soares and M.Fatima Bonaldo. "

BASE COUNT 118 a 84 c 110 g 118 t 2 others

ORIGIN

Query Match 100.0%; Score 15; DB 14; Length 432;

Best Local Similarity 100.0%; Pred. NO. 2.6e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCATGGTGCTCACT 15  
|||||  
Db 189 TCCATGGTGCTCACT 203

Search completed: September 11, 2003, 11:30:01  
Job time : 2295 secs